

Research and Development of High-End Computer Networks at GSFC

Presented at the Earth Science Technology Conference 2003
University of Maryland Conference Center
<http://esto.nasa.gov/conferences/esto2003/>

J. Patrick Gary
Network Projects Leader
Earth and Space Data Computing Division/Code 930
NASA Goddard Space Flight Center
pat.gary@nasa.gov
301-286-9539
June 25, 2003



Research and Development of High-End Computer Networks at GSFC

Outline

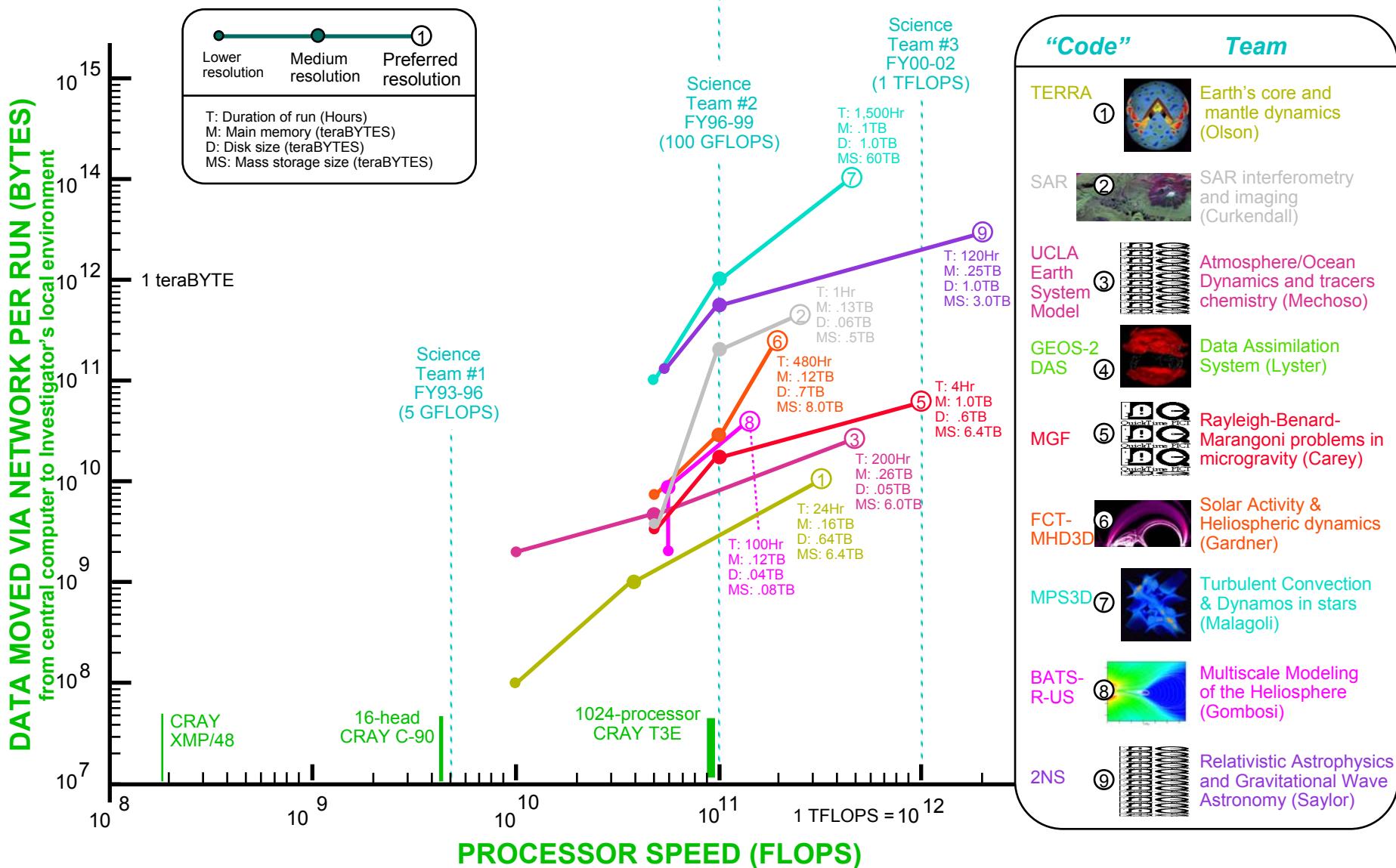
- Motivation
- Challenge
- nuttcp Performance Testing Tool
- Evaluations in Large Bandwidth*Delay Networks
 - » Examples of Pre-2000 Efforts
 - » Examples of Current Evaluations
- Conclusion

Research and Development of High-End Computer Networks at GSFC

Motivation

- Availability of new high-bandwidth networking technologies, e.g.:
 - » 1-10 Gigabit Ethernet switches/routers
 - » Wave division multiplexing on dark fiber
- New (and old) types of bandwidth-demanding applications, e.g.:
 - » Streaming HDTV and other real-time data over IP networks
 - » Storage Area Networks over IP networks
 - » High end computing

Long Haul Network vs CPU Requirements of ESS Grand Challenge Investigators (Science Team #2)



Research and Development of High-End Computer Networks at GSFC

Challenges

- Network users still do not always get the throughput performance that they should get or need to get
- Common factors affecting throughput performance
 - » Bandwidth or message size limitations in intermediate links of the end-to-end network path
 - » Limitations of the hardware/software network interfaces of the end user client workstations or servers
- Appropriate evaluation environment
 - » Large-scale fielding of advanced networking technologies
 - » Effective measurement and analysis tools, plus expertise to use
 - » Real applications/users willing to risk network testbed availability

Research and Development of High-End Computer Networks at GSFC

nuttcp Throughput Performance Analysis Tool

- Developed by GSFC's Bill Fink
- Determines raw TCP or UDP network layer throughput
 - » Transfers memory buffers from a source system across an interconnecting network to a destination system
 - » Transfers either a specified number of buffers or for a specified time interval, optionally with pacing or as multiple simultaneous streams
 - » Reports many statistics, including:
 - achieved network throughput in Mbps
 - user, system, and wall-clock time
 - transmitter and receiver CPU utilization
 - loss percentage (for UDP transfers)

Research and Development of High-End Computer Networks at GSFC

nuttcp Throughput Performance Analysis Tool (continued)

- Recognized in SC2002 tutorial as the recommended "great successor" to ttcp
- Already in extensive use in GSFC, DREN, Supernet, and MAX networks
- Example use:

```
[user@transmit_host]# nuttcp -t -T10 -w512 -b receive_host  
1183.125 MB / 10.01 sec = 991.7452 Mbps 82 %TX 37 %RX
```

- <ftp://ftp.lcp.nrl.navy.mil/u/bill/beta/nuttcp/>

→ GSFC <-> NRL OC-12 1 TB Challenge

shasta-a.nasa.atd.net



Sun UltraSPARC-2/300
Solaris 2.6 (128M)
SunATM-622 (2.1)

Using nttcp to transfer 1 TB of data
via Classical IP
(-l8192, -r134217728, -w512)

(5 h 52 m 26 s)
77% - 415.9686 Mbps - 85%

fozzie-a.lcp.nrl.navy.mil



Sun UltraSPARC-2/300
Solaris 2.6 (128M)
SunATM-622 (2.1)

For comparison purposes, at T1 speed,
it would take more than 66 days
to transfer 1 TB of data

HPCC ATM

ATDnet

NRL ATM

OC-12c ATM
MTU = 9180

Round Trip Time (RTT) ~ 1.4 ms
Maximum OC-12c ATM TCP Performance ~ 540 Mbps
Bandwidth*Delay ~ 92 KB

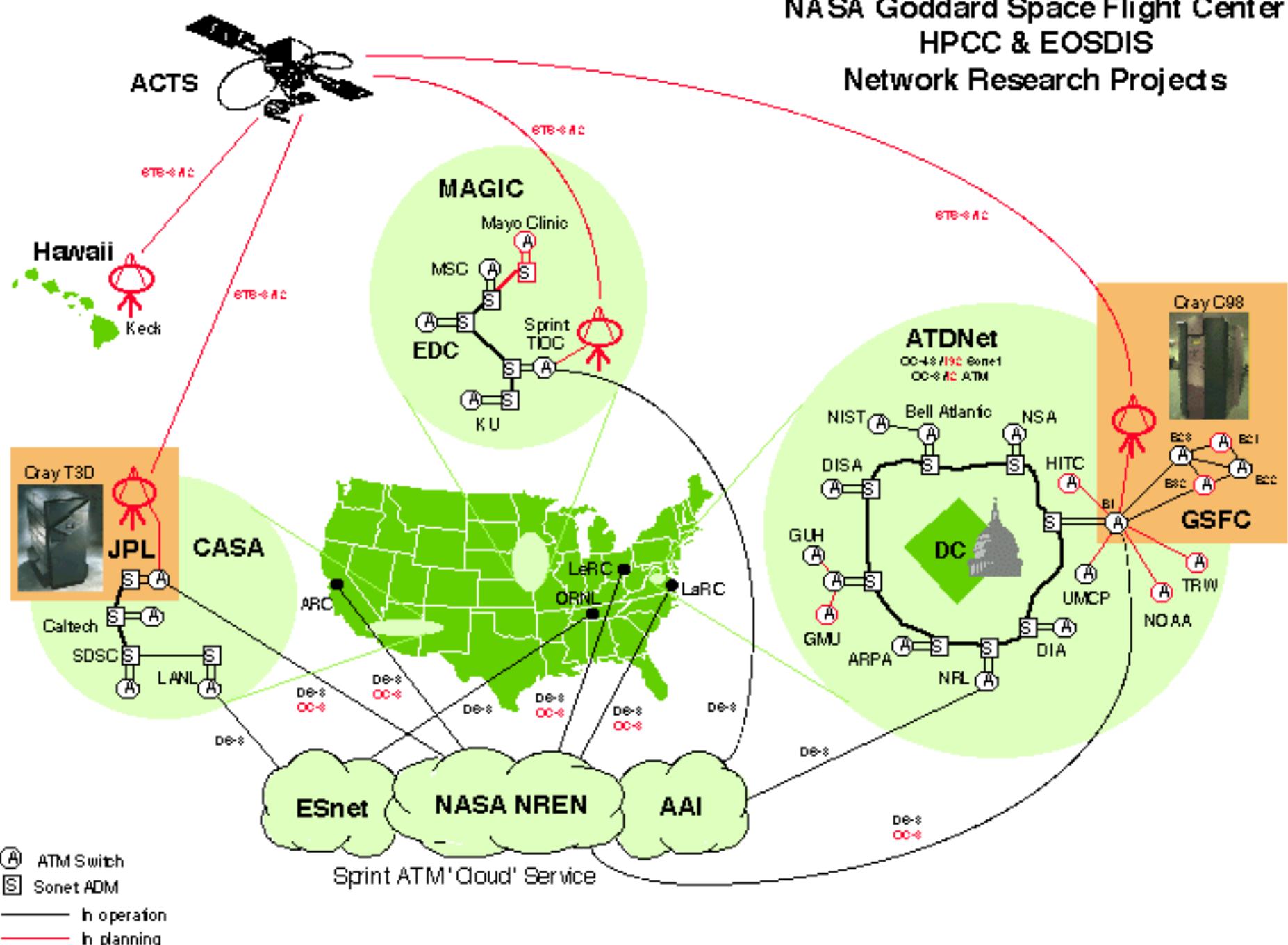
Research and Development of High-End Computer Networks at GSFC

Outline

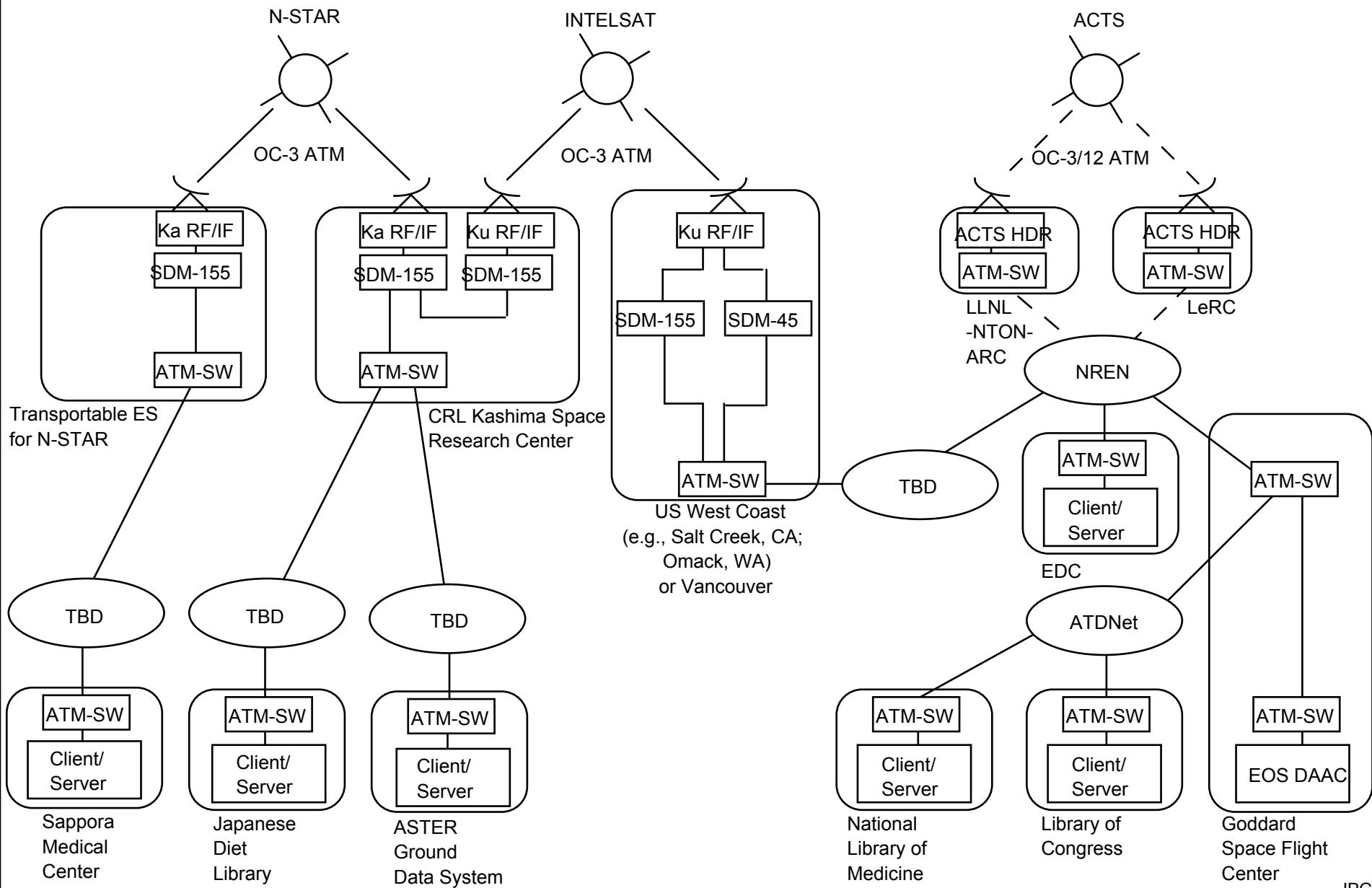
- Motivation
- Challenge
- nuttcp Performance Testing Tool
- Evaluations in Large Bandwidth*Delay Networks
 - » Examples of Pre-2000 Efforts
 - » Examples of Current Evaluations
- Conclusion



NASA Goddard Space Flight Center
HPCC & EOSDIS
Network Research Projects



Configuration of Networks for Trans-Pacific Digital Library Experiment



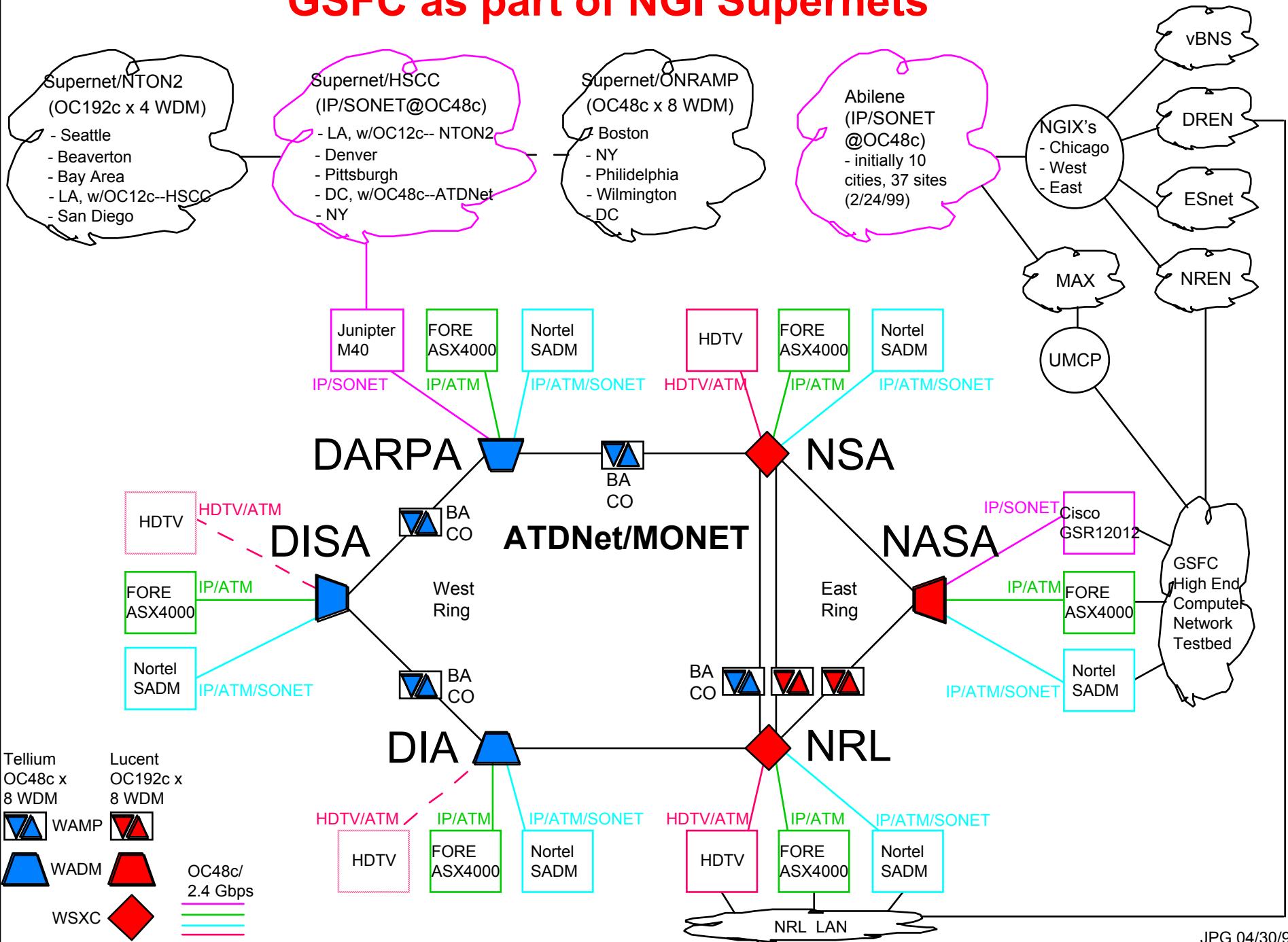
GSFC Benchmark Test Script & Key Findings in TPD

- Written to check and save information on the characteristics of the link prior to each Visible Human Viewer test run
- Test Script Checks
 - Roundtrip time (RTT) (using ping with small and large packet sizes)
 - Router hops (traceroute with small and large packets in both directions)
 - Transfer rates (ftp and nttcp of 7MB of data (size of largest image))

<u>Path</u>	<u>#</u>	<u>Path</u>	<u>Via</u>	<u>SkyX Proc</u>	<u>RTT (ms)</u>	<u>#Hops</u>	<u>ftp (Mbps)</u>	<u>nttcp (Mbps)</u>
					<u>65B/1500B</u>	<u>-> <-</u>	<u>15KB/7MB</u>	<u>7MB</u>
1	SMU-GSFC	Intelsat		Yes	1124/1127	14/14	/15.2	11.9
2	SMU-NLM	Intelsat		Yes	1127/1130	16/16	10.9/15.2	11.9
3	SMU-NLM	Intelsat		No	1127/1130	16/16	.026/.224	0.225
4	SMU-GSFC	TransPAC		No	191/224	16/14	/.817	0.732

where Intelsat is the satellite path and TransPAC is the terrestrial path

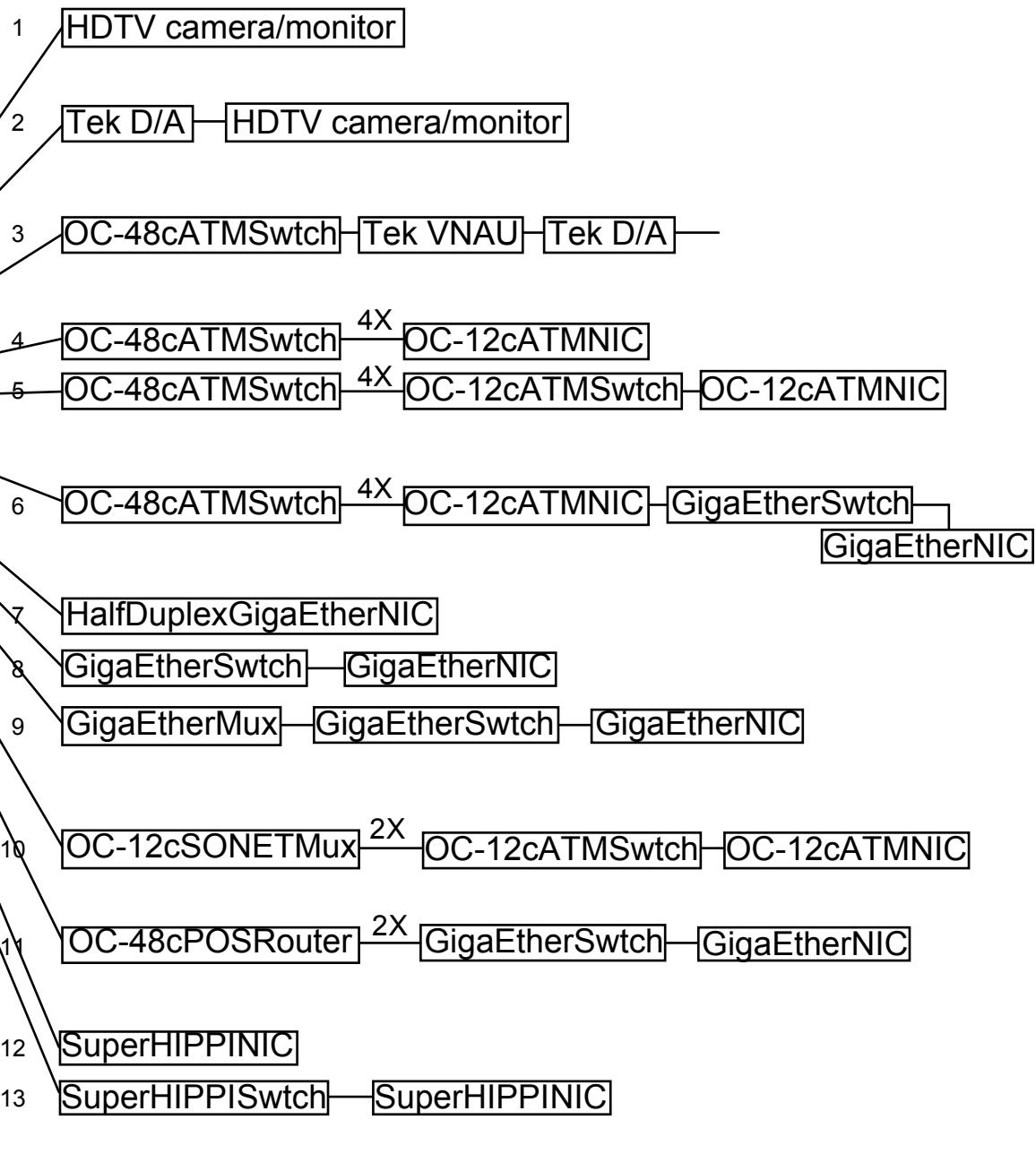
GSFC as part of NGI Supernets



Different Protocol Stacks/Traffic Types in ATDNet/MONET



At least two 8 x OC-48c WDM MAN links and eight OC-48c local ports at each of DARPA, DIA, DISA, NASA, NRL, and NSA



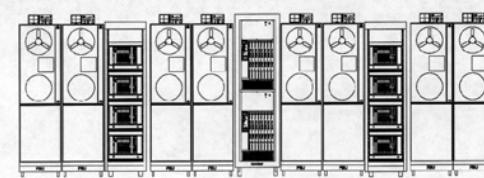
Schematic of Gbps e-VLBI Demonstration Experiment



Westford

~1.5 km

Mark 4
Correlator



Haystack Observatory

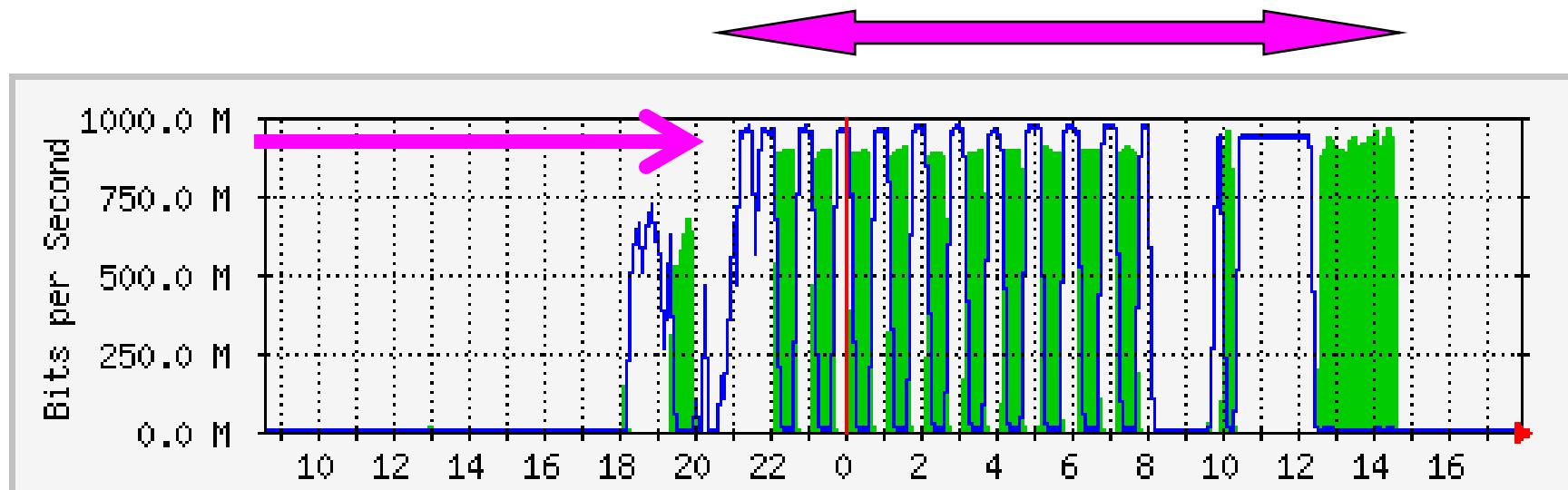
~650 km



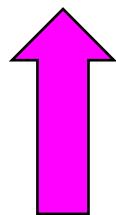
NASA/GSFC

**Glownet, Bossnet,
MAX, NASA/HECN
network segments**

e-VLBI GGAO-Haystack Data Rates Sustained During a 16-Hour-Long Evaluation Test



Max In:970.5 Mb/s (97.1%) Average In:210.8 Mb/s (21.1%) Current In:168.0 b/s (0.0%)
Max Out:978.1 Mb/s (97.8%) Average Out:263.6 Mb/s (26.4%) Current Out:216.0 b/s (0.0%)



Ethernet Jumbo Frame versus Standard Frame: Effect on Data Access Performance (Tests Performed by GSFC's Bill Fink, 10/30/02)

- Test configuration
 - » Data Server: Maximum Throughput Sledgehammer SH 200 network attached storage
 - » Data Client: 867 MHz Macintosh G4
 - » Access Protocol: NFS v3.0
 - » Interconnection Network: Extreme Network Summit 51 Gigabit Ethernet switch (includes jumbo frame capability)
- Performance (MegaBytes per second) with jumbo frames (MTU=9000, NFS rsize=8192,wsize=8192)

<-----Transmit----->			<-----Receive----->		
» Min	Avg	Max	Min	Avg	Max
» 38.1592	38.4785	38.8381	36.0036	44.9774	52.7600
- Performance (MegaBytes per second) with standard frames (MTU=1500, NFS rsize=1024,wsize=1024)

<-----Transmit----->			<-----Receive----->		
» Min	Avg	Max	Min	Avg	Max
» 4.6198	4.6387	4.6593	4.6263	4.6392	4.6456

Research and Development of High-End Computer Networks at GSFC

Conclusion

- Inexpensive 1 and 10 Gbps GigE and WDM optical networking technologies can significantly enable Earth science applications
- But knowledgeable selection and use of those technologies can only be achieved through advanced network technology evaluations
- GSFC's HECN Project continues to be every successful at these advanced network technology evaluations

Research and Development of High-End Computer Networks at GSFC

Acknowledgments

- Jim Fischer (GSFC), ESTO/CT Project Manager
- Bill Fink (GSFC), developer of nuttcp and technical lead on most HECN advanced technology evaluations
- Rest of the present HECN Team: Herb Durbeck (GSFC), Kevin Kranacs (GSFC), Lee Foster (GSFC), Paul Lang (ADNET), Aruna Muppalla (ADNET), Wei-Li Liu (ADNET), and Chandu Rathod (ADNET)
- Former HECN Team members: Kalyan Kidambi, Marian Stagarescu, and Sanjiv Duggal (all then RITSS)
- Jeff Martz (CSC), wiring and network equipment installation super-expert

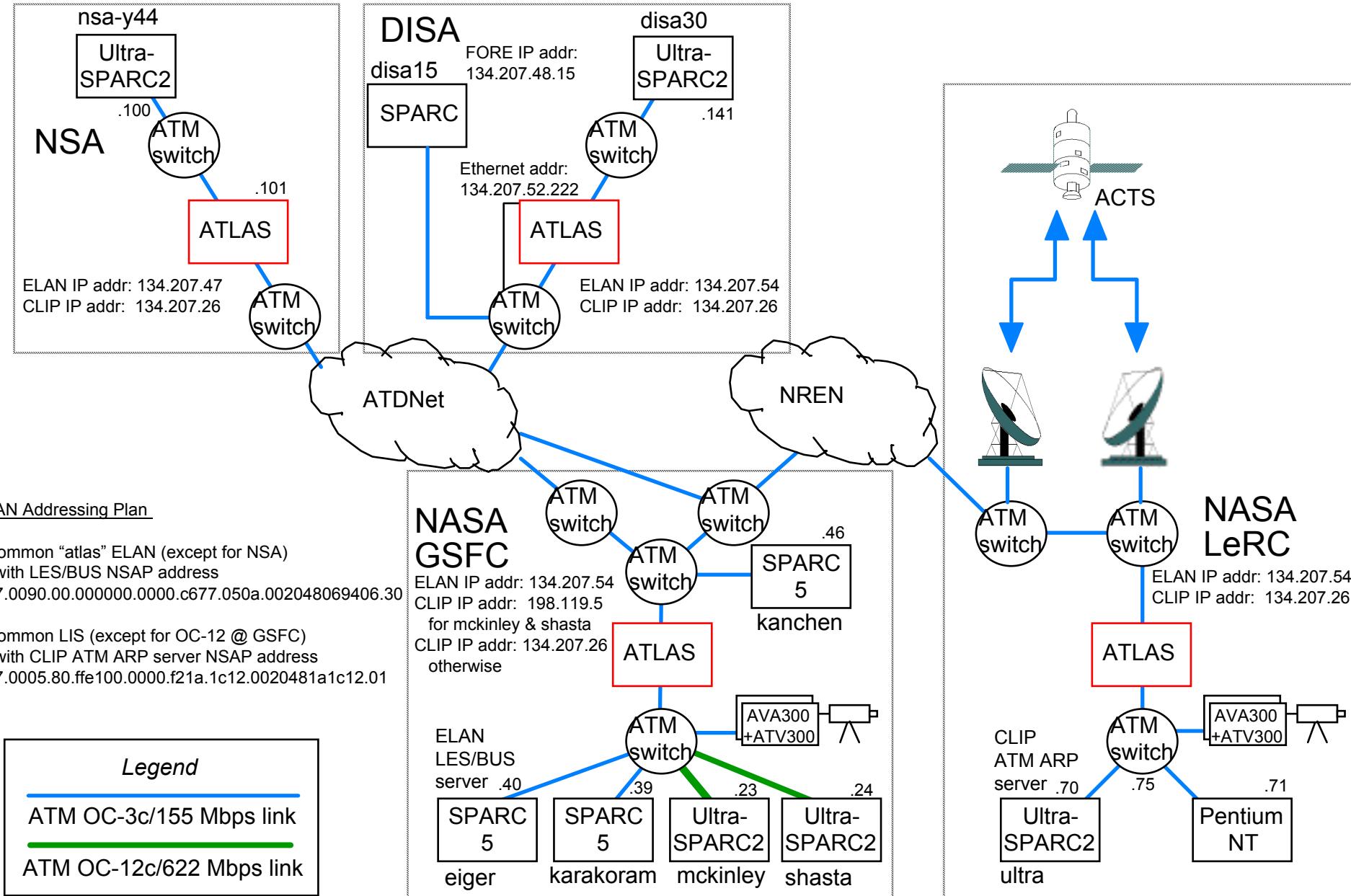
Research and Development of High-End Computer Networks at GSFC

Backup Charts

Usage (transmitter): nuttcp -t [-options] host [<in >]

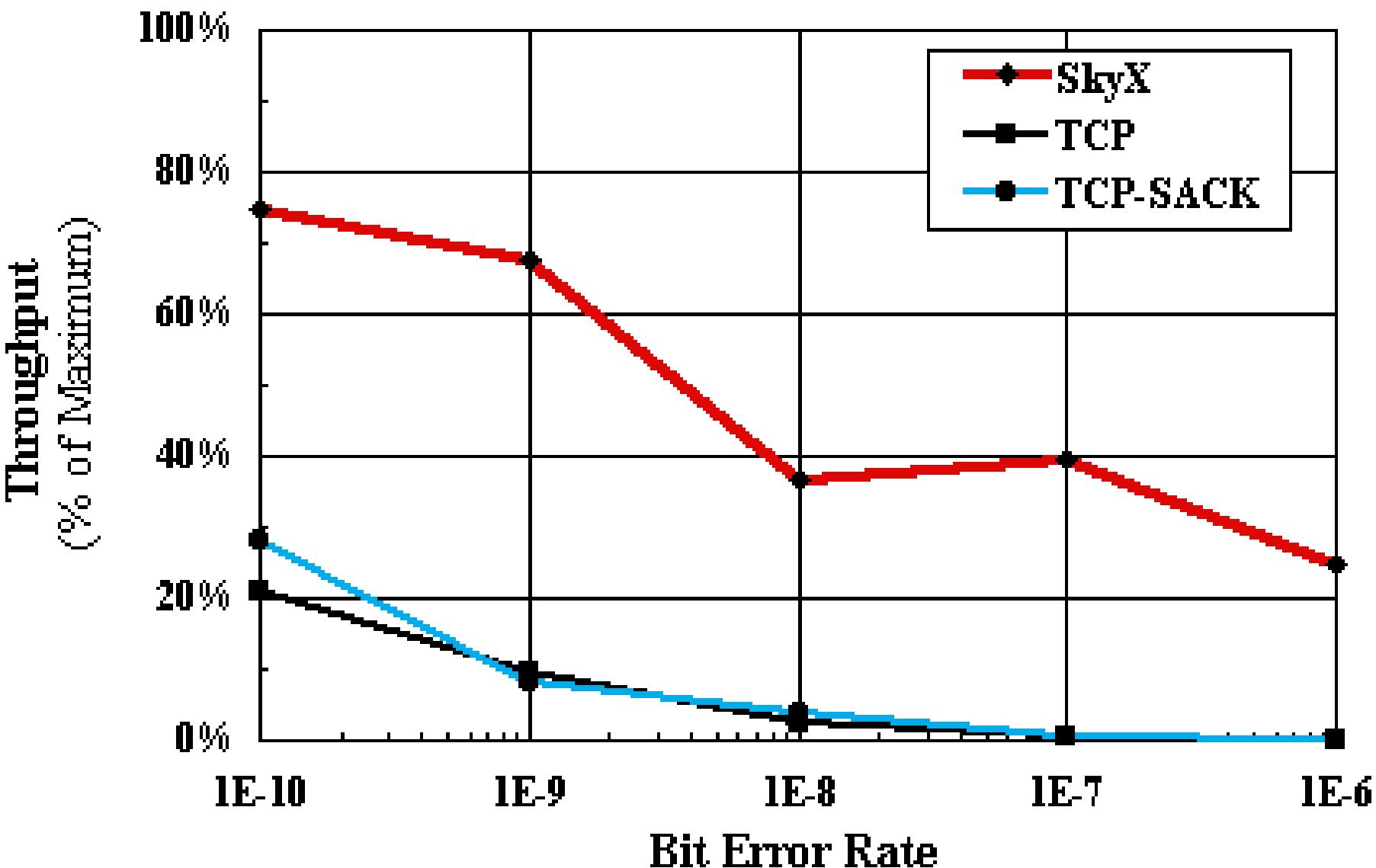
- l## length of network write buf (default 8192/udp, 65536/tcp)
- s don't source a pattern to network, use stdin
- n## number of source bufs written to network (default 2048)
- w## transmitter window size in KB
- ws## server receive window size in KB
- wb braindead Solaris 2.8 (sets both xmit and rcv windows)
- p## port number to send to (default 5001)
- P## port number for control connection (default 5000)
- u use UDP instead of TCP
- D don't buffer TCP writes (sets TCP_NODELAY socket option)
- N## number of streams (starting at port number)
- R## transmit rate limit in Kbps (or (m|M)bps or (g|G)bps)
- T## transmit timeout interval in seconds (or (m|M)inutes)
- i## server interval reporting in seconds (or (m|M)inutes)
- lxxx identifier for nuttcp output (max of 40 characters)
- F flip option to reverse direction of data connection open
- xP## set nuttcp process priority (must be root)
- d set TCP SO_DEBUG option on data socket
- v verbose output
- b brief output

Configuration for SPOCK Evaluation of STK/NSG ATLAS ATM Firewall



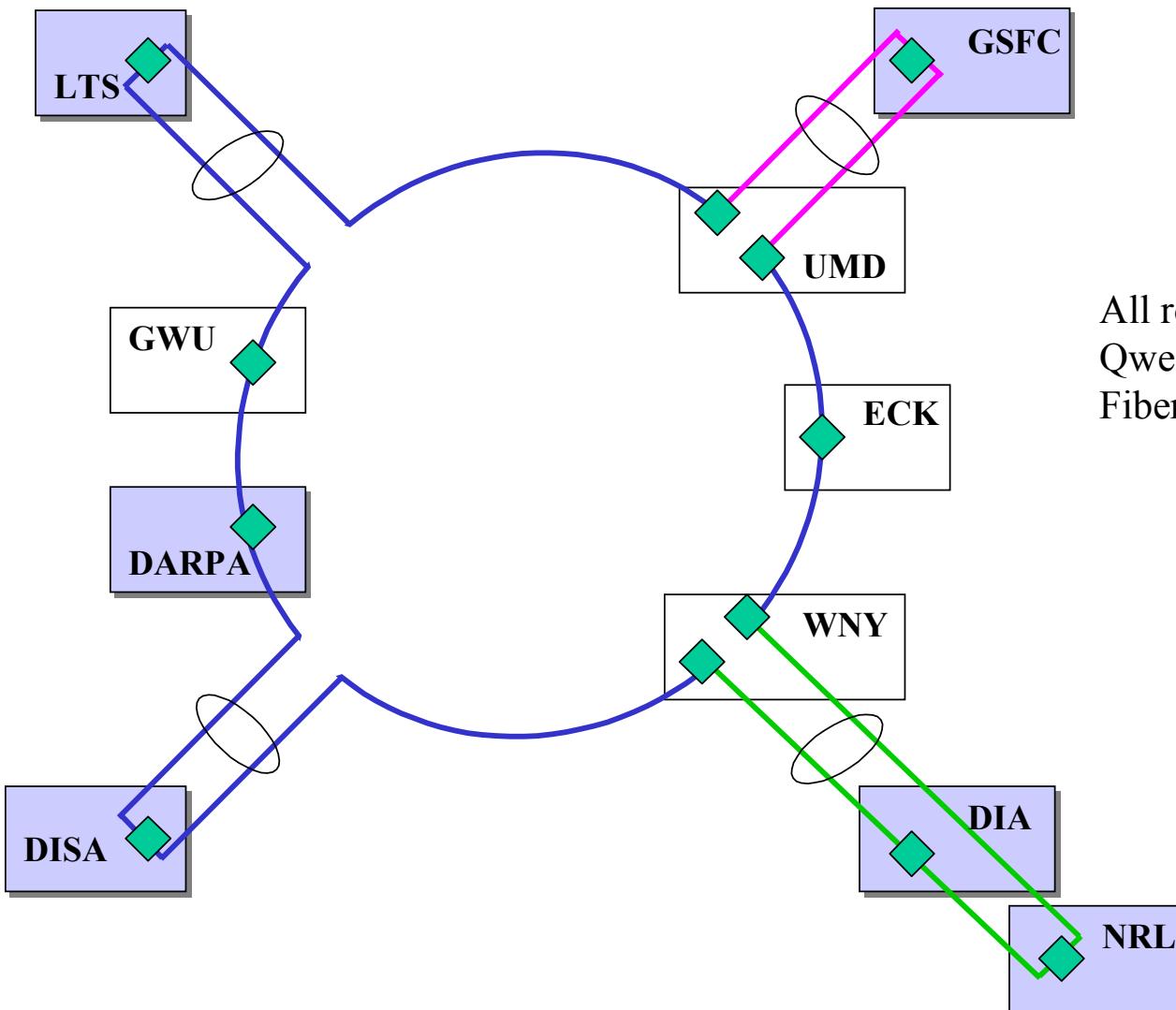
SkyX and TCP Throughput vs. BER

Satellite Conditions: RTT = 540 ms



ATDnet(V2) Ring Topology

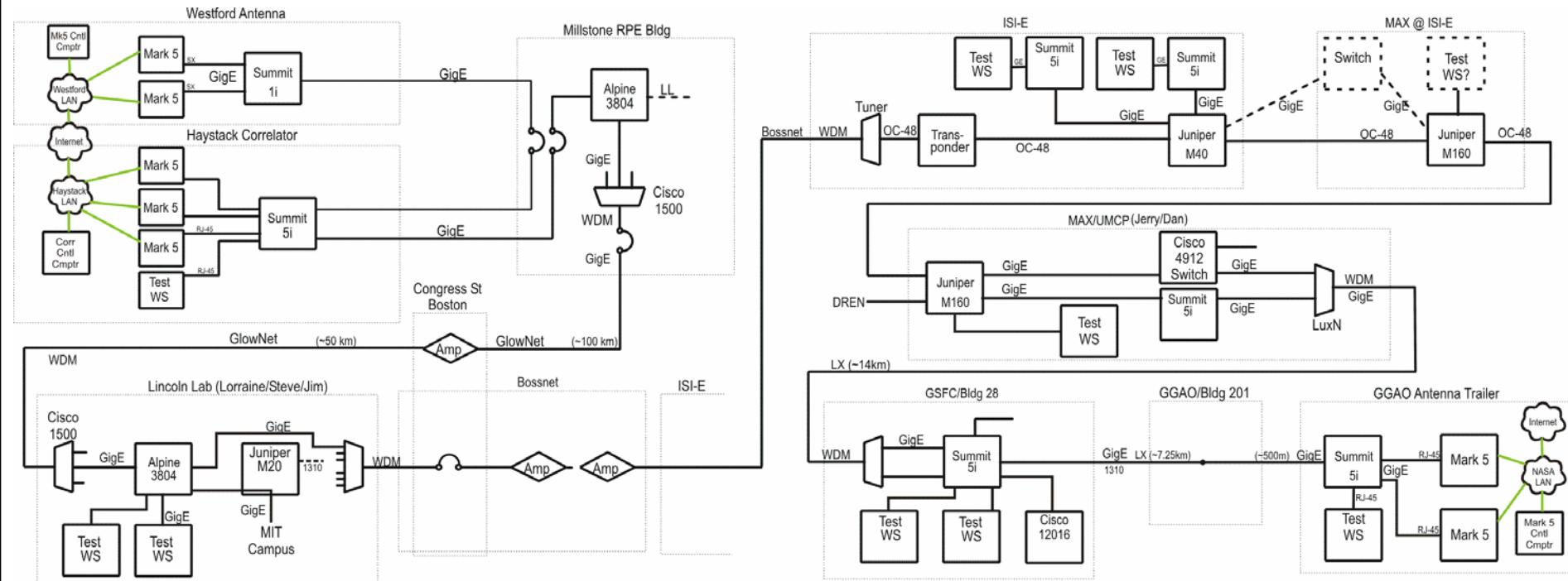
Basic configuration using Qwest to reach LTS



All routes are 2 strands fiber
Qwest routes are Lucent Truewave
Fibergate routes are SMF28

- Fiber bundle
- Cross-connect
- Primary site
- Ancillary site
- Qwest fiber
- Fibergate fiber
- Govt supplied fiber

Details of the e-VLBI Network Path



For more info see ftp://web.haystack.edu/pub/e-vlbi/demo_report.pdf.

GSFC SANoverIP iSCSI Eval Configuration

